DOCUMENT RESUME

ED 403 302 TM 026 026

AUTHOR Davidson, Bruce W.; Dunham, Rodney L.
TITLE Assessing EFL Student Progress in Critical Thinking with the Ennis-Weir Critical Thinking Essay Test.
PUB DATE 96
NOTE 22p.; Paper presented at the Annual International Conference of the Japan Association for Language Teaching (21st, Nagoya, Japan, November 5, 1995) and at the International Conference on Critical Thinking and Educational Reform (16th, Rohnert Park, CA, July 18-31, 1996).
PUB TYPE Reports - Research/Technical (143) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *College Students; *Critical Thinking; *English (Second Language); Essay Tests; Females; Foreign Countries; *Second Language Learning; *Teaching Methods; Two Year Colleges
IDENTIFIERS *Ennis Weir Critical Thinking Essay Test; Japanese People

ABSTRACT

Recent trends in the teaching of English as a Foreign Language (EFL) or English as a Second Language (ESL) have emphasized the importance of promoting thinking as an integral part of English language pedagogy; however, empirical research has not established that training in thinking skills can be combined effectively with EFL/ESL instruction. In this study, the Ennis-Weir Critical Thinking Essay Test was used to assess progress in critical thinking after a year of intensive academic English instruction for 36 Japanese students enrolled in a private two-year women's junior college in Oaska, Japan. A control group received only content-based intensive English instruction, while the treatment group received additional training in critical thinking. The treatment group scored significantly higher on the test ("p" =0.000). The results imply that critical thinking skills can indeed be taught as part of academic EFL/ESL instruction. (Contains 3 tables and 26 references.) (Author/SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

Running Head: Assessing EFL Student Critical Thinking

ED 403 302

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY
BRUCE W. DAVIDSON
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

Assessing EFL Student Progress in Critical Thinking With the Ennis-Weir Critical Thinking Essay Test

Bruce W. Davidson
Hokusei Gakuen University
Rodney L. Dunham
Tezukayama College

TM 026026

BEST COPY AVAILABLE

Abstract

Recent trends in EFL/ESL have emphasized the importance of promoting thinking as an integral part of English language pedagogy; however, empirical research has not established that training in thinking skills can be effectively combined with EFL/ESL instruction. This study made use of the Ennis-Weir Critical Thinking Essay Test to assess progress in critical thinking after a year of intensive academic English instruction among Japanese students ($N = 36$). A control group received only content-based intensive English instruction, while a treatment group received additional training in critical thinking. The treatment group scored significantly higher on the test ($p = .000$). The results imply that critical thinking skills can indeed be taught as part of academic EFL/ESL instruction.

Assessing EFL Student Progress in Critical Thinking With the Ennis-Weir Critical Thinking Essay Test

Since the advent of research into cognitive development, language teachers and linguists generally recognize the close connection between language learning and thinking processes. In particular, ESL reading research has shown some correlation between ESL reading comprehension and familiarity with the formal or content schemata of English texts (Carrell, 1987). Furthermore, noting the unreflective character of many language-teaching approaches that only encourage verbal output or passive input, Tarvin and Al-arishi (1991, 1994) have explored some methods to make language teaching more thoughtful. Similarly, Chamot (1995) has argued from current educational trends promoting higher-order thinking that EFL/ESL teachers also need to turn the classroom into a "community of thinkers." Informal observations may indicate that thinking skills can indeed be taught in an EFL/ESL context (Davidson, 1994, 1995). Without formal testing, however, it is difficult to establish concretely that they can be. Though there has been a lot of thought and research devoted to the development of critical thinking skills in native English speaker educational programs of various sorts, there has been little research in the area of combining critical thinking with EFL/ESL instruction.

Recently content-based intensive English instruction has also proven to have many advantages and possibilities (Snow & Brinton, 1988). Is one of them the promotion of critical thinking skills through thought-provoking content? It might be expected that such abilities will develop through discussion, reading, or composition about subjects requiring some serious analytical attention; however, Chance's (1986) survey concluded that critical thinking skills do not develop simply as a by-product of the study of specific subjects. In addition, Halpern (1993) cites evidence from various sources that critical thinking skills can be inculcated through explicit instruction.

These issues and findings inspired a pilot study to discover whether or not critical thinking could be taught to Japanese students of English in a content-based EFL program. After defining what we mean by "critical thinking," we will describe the

intensive English program the subjects were enrolled in along with the specifics of the current study. Two research questions guided us:

1. On a critical thinking test task, will English learners exposed to critical thinking skills-training do significantly better than similar students who have not received such training?
2. Can a critical thinking test designed for native English speakers be used as an instrument for evaluating critical thinking skills among non-native English learners?

The Critical Thinking Concept and An Inventory of Component Skills

Critical thinking involves rational judgment and discernment of the elements of reasoning, and various definitions of critical thinking reflect this. Norris and Ennis (1989) explain critical thinking as "reasonable and reflective thinking that is focused upon deciding what to believe and do" (p. 3), a definition also stated somewhat differently by Lippman (1991), who defines it as the inculcation of healthy skepticism, and Siegel (1988), who considers the critical thinker to be one who "is appropriately moved by reasons" (p. 2). In contrast to rote memorization or simple information recall, methods for encouraging critical thinking have as their goal the stimulation of the analytical and evaluative processes of the mind (Paul, 1992). Ennis and Norris (1989) have listed a number of critical thinking abilities to develop, which they group in the following manner:

Elementary Clarification

1. Focusing on a question
2. Analyzing arguments
3. Asking and answering questions that clarify and challenge

Basic Support

4. Judging the credibility of a source
5. Making and judging observations

Inference

6. Making and judging deductions
7. Making and judging inductions
8. Making and judging value judgments

Advanced Clarification

9. Defining terms and judging definitions

10. Identifying assumptions

Strategies and Tactics

11. Deciding on an action

12. Interacting with others (p. 14)

This concept of critical thinking and inventory of skills inspired both the instructional treatment and the selection of the Ennis-Weir test in this study.

Method: Subjects and Treatment

All the participants in the study were first-year students enrolled in a private two-year women's junior college in Osaka, Japan. The college's curriculum consisted mainly of an intensive academic English program. Weekly English courses included Oral Discussion (3 hours), Composition (2 hours), Reading (3 hours), Pronunciation (3 hours), and Grammar/Listening (2 hours), totaling 13 hours a week, considering each fifty-minute class session as an hour. Oral Discussion, Reading, and Composition followed a topical syllabus, which included such themes as Prejudice/ Human Rights, Advertising/Consumerism, and Women's Issues/Child-raising. Along with the topic, each unit also introduced a rhetorical mode: Illustration, Process, Definition, Classification, Comparison/Contrast, and Persuasion. The first three composition units required students to write a paragraph using each mode, and the last three units progressed to multi-paragraph essays. The Unit 6 persuasion essay was written in a mini-term paper format, with references. In addition, students took a weekly one-hour seminar course, which was conducted in English and concerned some interesting topic or theme. The treatment group was composed of students from a seminar on Critical Thinking. The integrated, content-based aspect of the program is meant to involve students in in-depth analysis and expression concerning subjects significant in their own lives and in Japanese society. This course of study would seem well-suited to encouraging the development of critical thinking skills as a by-product, since the topics all necessitate thought.

Students in the study had varying degrees of English proficiency as measured by an in-house proficiency test. At the beginning of the year this test divided students into five levels of classes according to their scores: A, B, C, D, and E. The A classes had the highest level of proficiency and included many students returning from a year or longer of study abroad, whereas the E classes were much less proficient and included some students who had entered solely by high school recommendation, never having taken the English entrance examination normally given to prospective students. Regardless of proficiency, however, all classes received similar instruction based on the same content and rhetorical modes noted previously. 19 volunteers not enrolled in the critical thinking seminar course served as a control group. These 19 students represented a broad range of English proficiency levels, as measured by the in-house test, and the group was comparable in that respect to the treatment group, which contained a similar range. The treatment group consisted of 17 out of 22 members of the seminar on critical thinking. 5 could not take the Ennis-Weir test due to circumstances arising from the Kobe-Osaka earthquake. Since students enrolled in seminars through a semi-lottery system, the authors consider that in this case group assignment generally embodied the spirit of randomness, although it was not completely random. No pre-test was given, in line with the advice of Ennis and Weir (1985), who state that a pre-test is not necessary in research using the test as long as a control group exists. Babbie (1983) has noted that a posttest-only control group design is quite acceptable as long as group assignment is random.

The treatment group took part in a course designed to train them in some basic elements of critical thinking: source credibility, inductive reasoning, informal deductive logic, and assumption identification. These broad categories encompass most of Ennis's and Norris's (1989) list of critical thinking skills, so they were adopted as a framework for the seminar course. During the first semester, instruction dealt with inductive reasoning and source credibility; during the second semester, the emphasis was deductive reasoning and assumption-identification. The course began with an introduction to the concept of critical thinking. Sessions were devoted to exploring various kinds of reasoning fallacies and misuse of evidence, such as over-generalization and the false

dilemma (Chaffee, 1991; Damer, 1995). Students were given lists of brief fallacious arguments and asked to explain the problems of each in their own words. In the second half of the semester, the focus shifted to source credibility. Students did exercises in which they evaluated varying accounts of the same event according to differing viewpoints. For example, in groups they discussed and ranked varying accounts of the results of an international conference, keeping in mind a list of question-criteria: Does the news presenter have a reason to be biased? Is the source an expert in the field?, etc. (Beyer, 1991). As homework students brought in similar examples to present and evaluate. In the second semester the emphasis shifted to basic argument analysis. First, students did exercises to help them distinguish real arguments from bare claims offering no reason (Engel, 1994). Then they identified the claims and supporting reasons contained in brief arguments. Later in the semester the instructor introduced less-obvious aspects of deductive reasoning: unstated assumptions and implications (Scriven, 1976). Using magazine advertisements and other material, students practiced identifying assumptions and implications. As a result of the earthquake and other circumstances, a total of only 18 class hours was actually devoted to teaching course content.

The Ennis-Weir Critical Thinking Essay Test

Test Description

By means of the Ennis-Weir test, the researchers hoped that real progress in thinking skill might be concretely confirmed in the case of students exposed to skills training in critical thinking. This test was chosen for various reasons. One is that it is one of the most generally well-accepted measuring instruments among educators in the critical thinking movement (Walsh & Paul, n.d.), and inter-rater reliabilities have been very high when it has been used (Ennis & Weir, 1985; Hatcher, 1995). Another is that, in contrast to multiple-choice tests, the test allows students to justify varying responses, and the test itself presents a realistic critical evaluation task. Other critical thinking tests are available, but almost all of these are multiple-choice instruments that suffer from various weaknesses, such as background bias and the impossibility of knowing the reasoning behind an examinee's answer-choice (Ennis, Millman, & Tomko, 1985; Norris & Ennis,

1989). Furthermore, the relatively simple subject-matter and language of the Ennis-Weir test make it suitable for non-native speakers. It has been used successfully with first-year junior high school native English speakers in the U.S. The test itself contains a simple set of instructions and a letter to a newspaper editor containing ten brief paragraphs. The fictional writer, Raywift, recommends that overnight parking be prohibited on all the streets of his town, Moorburg. After a brief introduction, eight numbered paragraphs elaborate the argument. Most are weak and commit various common reasoning fallacies such as equivocation, irrelevancy, poor statistical sampling, and circular reasoning, but some contain legitimate support, consisting in the use of qualified experts or a relevant reason. Point-by-point, the examinee's task is to judge the thinking of each paragraph and to evaluate the strength of the letter's argument as a whole in a final summary-paragraph.

(TABLE 1 HERE)

To give readers some idea of the contents of Raywift's letter, here we will include verbatim two of the numbered paragraphs, the third and the sixth:

> 3. Traffic on some streets is also bad in the morning when factory workers are on their way to the 6 a.m. shift. If there were no cars parked on these streets between 2 a.m. and 6 a.m., then there would be more room for this traffic.
>
> 6. Last month, the Chief of Police, Burgess Jones, ran an experiment which proves that parking should be prohibited from 2 a.m. to 6 a.m. On one of our busiest streets, Marquand Avenue, he placed experimental signs for one day. The signs prohibited parking from 2 a.m. to 6 a.m. During the four-hour period, there was <u>not one accident</u> [italics added] on Marquand. Everyone knows, of course, that there have been over four hundred accidents on Marquand during the past year (Ennis & Weir, 1985, p. 13).

In responding to Paragraph 3, examinees are expected to notice that a relevant reason is offered to support Raywift's argument. Similarly, test-takers are supposed to show some indication that they comprehend the obvious flaws of the experiment in Paragraph 6.

A very clear and specific scoring protocol accompanies the test indicating various possible answers and how each is to be scored. Points are awarded both for judging correctly and for indicating a valid reason for one's judgment, and a penalty point of -1 can be deducted for poor reasoning. A range of possible correct answers is given. Each answer can receive a maximum of 3 points and a minimum of -1, except for the summary paragraph nine, where a maximum of 5 points can be awarded. Therefore, the overall score can range from -9 to +29. In general, the protocol gives latitude to raters to award points whenever an examinee can give a credible reason in support of his or her evaluative judgment, even when his or her judgment differs from that of the protocol writers. Brief answers are acceptable as long as they indicate a valid judgment, backed up with a sound reason for that judgment.

A limited amount of research has been done in the U.S. using the test. The largest study to date has been Hatcher's (1995) at Baker University. Over a period of four years (1990-1994), Baker University American freshmen scored an average of 11.8 to 13.8 on the Ennis-Weir test after a year-long compulsory critical thinking course. They had scored from 5.8 to 9.4 on a pretest and registered gains of 2.8, 5.8, 5.8, and 6.0 points. Interestingly, a number of Chinese and Japanese students at Baker University also took part in the study, but their scores were eliminated from it because they scored consistently very poorly on the test. Hatcher (1994) speculated that their low scores may be due to Oriental politeness and a hesitancy to criticize the Moorburg letter. We hoped not to have the same difficulty in making use of the test.

Test Administration

In the last week of second semester classes, the Ennis-Weir Critical Thinking Essay Test (1985) was administered, and the control group was given the test within the same week. Both groups were given eighty minutes to read the test and write nine brief paragraphs in response to the Moorburg letter. This is twice the amount of time recommended by the test-makers. Since the subjects were non-native English students, it was felt that much more time would be necessary for them to comprehend the material and compose answers. To help them with the language aspects of the test, they were

allowed to use dictionaries. Furthermore, before taking the test, all subjects received two sample test items with model answers to make sure that students understood two things: (1) that they had to make a clear evaluative judgment as to whether the argument in each paragraph was a good one or not and (2) that they had to give a clear reason or explanation for their judgment. Without such explicit direction, the subjects might not have done either of these two things. However, students in the present study were used to doing peer-evaluation of essays in composition classes, so the idea of writing comments or criticism about a piece of writing was already somewhat familiar to them.

Results

Tests were scored blindly and independently by two raters. The test-raters in this study found little difficulty in using the protocol to judge student answers. Grammatical or vocabulary problems were overlooked unless they made an answer incomprehensible. The scores and some basic student information collected on the answer form, such as English level and travel experience, were entered into a Macintosh LC 630 computer using the SPSS program version 4.0. Inter-rater reliability was found to be high (r = .72). Therefore, we used an SPSS compute statement to average each of the ten scores (one for each of the nine paragraphs and a total score) given by the two raters for each student. The computed scores were used for all subsequent analysis.

The small number of students in the study (control n = 19 and treatment n = 17) makes relationship detection difficult unless there is a very strong relationship. Because this is the first study of this type, we were interested in detecting moderate relationships as well as strong ones. Consequently, we decided that the risk of committing a Type I error would be less important than missing moderate relationships. Therefore, the significance level of .10 was chosen as the cutoff for accepting or rejecting relationships. Nevertheless, we will report here the exact probability for all statistical results that indicated significant relationships.

The most important analysis, of course, dealt with the effect of critical thinking training on test scores. The mean for the treatment group's score on the Ennis-Weir Test was 6.6, which was significantly higher than the control group's mean score of 0.6 (t

(27.73) = -4.99, p = .000). Table 2 shows the range of scores for each group and details the differences in the scores. As the table shows, 10 students in the treatment group scored 7 or higher, while only 1 scored 0 or lower. In contrast, in the control group, 9 students scored 0 or lower, and no score was higher than 6.5.

(TABLE 2 HERE)

The next analysis compared the individual paragraph scores of the control and treatment groups. There were significant differences between the scores of the control and the treatment groups on two paragraphs: the third paragraph (M = -0.18 and 0.50, respectively, t (33.87) = -2.10, p = .043) and the sixth paragraph (M = -0.55 and 0.15, respectively, t (27.10) = -2.59, p = .015). The difference in scores on the eighth paragraph approached significance (t (33.69) = -1.71, p = .096), with the treatment group scoring higher (M = 0.47) than the control group (M = -0.52).However, the scores on the remaining paragraphs showed no statistically significant difference between the control and treatment groups (p > .10).

Since the test was in English, a foreign language for the subjects, proficiency may have affected their scores. Except for proficiency level C (n = 14), the number of students in each level was quite small (A level n = 7, B level n = 5, D level n = 5, and E level n = 5). Therefore, a new, three-level variable was created using the SPSS "if" command to combine A students with B students and to combine D students with E students, while leaving C intact. The crosstabs command of the SPSS program produced a table with a fairly even distribution of students among the three English levels in the control group (n = 5, 10, and 4, respectively) and the treatment group (n = 7, 4, and 6, respectively). This distribution indicates that there was no relationship between English proficiency and the type of group (X^2 (2, N = 36) = 3.204, p = .202). An analysis of variance was also run to examine the relationship between English level and scores on the test. There was no significant relationship between the two variables (F (2, 35) = 1.57, p = .224). Table 3 shows that the range of scores was comparable for each proficiency level. Judging by this analysis as well as the phrasing of student answers on the test, we believe that students generally did not do poorly simply as a result of an inability to

understand the contents of the test. The sample test items appeared to succeed in helping students to grasp the kind of test task they were engaged in, and the wording of the test did not appear to present an insurmountable problem even to lower-level students.

(TABLE 3 HERE)

The next variable examined was overseas experience, since a number of students had lived a year or longer in an English-speaking country. Using a T-test, scores of students who had traveled overseas and those who had not were compared. The differences in scores between the two groups was not significant for total scores or for any individual paragraph score ($p > .10$).

Because the Ennis-Weir test deals with parking problems, each student was asked to report on the test form whether or not she possessed a driver's license. Japanese students often first learn to drive at the age of nineteen or twenty, the age of the young women in this study, and familiarity with driving an automobile may have helped some students do better on the test, which concerns a parking problem. Scores for the two groups (those with licenses and those without) were compared. Total scores were statistically the same for both groups; however, students without driver's licenses ($\underline{M}$ = -0.66) scored significantly lower on the seventh paragraph than students with driver's licenses ($\underline{M}$= -0.17) ($\underline{t}$ (21.14) = -1.84, $\underline{p}$ = .079). On the eighth paragraph, students without driver's licenses ($\underline{M}$ = -0.24) also scored significantly lower than students with driver's licenses ($\underline{M}$ = 0.70) ($\underline{t}$ (31.50) = -3.31, p = .002). Otherwise, the fact of having a driver's license had no significant relationship with student scores ($p > .10$). Since the specific issues addressed in the seventh and eighth paragraphs are not directly related to the experience of driving, we consider the statistical significance to be unrelated to the current study.

To sum up, these statistical analyses appear to indicate that the differences in scores between the treatment and control groups cannot be accounted for by differences in English proficiency levels or other factors such as overseas experience or having a driver's license. Therefore, the differences in scores on the Ennis-Weir Test can probably be attributed to the critical thinking training given to the treatment group.

Discussion and Conclusion

Both research questions that guided us can be answered in the affirmative, based on the results of this study. The Ennis-Weir test, designed for native English speakers, appears to be usable in the case of non-native English learners. Furthermore, it is encouraging to find that even a small amount of instruction in the basics of critical thinking appeared to result in higher scores in the treatment group. Critical thinking skills can apparently be taught to some extent along with English as a foreign language and can therefore enhance a content-based course of study. In view of the relatively small amount of actual instruction, the rather low average score of 6.6 is not surprising, and it is much better than the performance of the control group ($\underline{M} = 0.6$). As a point of comparison, Baker University American freshmen registered gains of 2.8, 5.8, 5.8, and 6.0 (Hatcher, 1995) in four successive years. Interestingly, three of those gains approximate the difference of 6.0 that we found in the mean scores of our two groups, though the mean score of the treatment group (6.6) is only half that of the average post-test scores of the Baker freshmen. Looking at the individual test items, differences between the two groups appeared specifically in items which had received instructional attention in the critical thinking class. Paragraph 6 deals with the misuse of statistics, a reasoning problem dealt with in class, while Paragraph 3 featured a relevant reason, another instructional point. The difference in performance on Paragraph 8, which concerned the use of experts and their credibility as sources, also approached statistical significance, and that area also had received attention in the source-credibility component of the seminar. In contrast, little difference in scores appeared in the case of Paragraphs 1 and 7, which both concerned inappropriate definitions, an area not dealt with in the critical thinking course. Furthermore, there was little difference in scores on Paragraph 4, which consists in circular reasoning and is very similar to one of the sample test items. Perhaps because of its similarity, 35 students responded correctly to it.

The overall quality of the answers of the two groups differed, but they shared certain tendencies indicating a general weakness in the area of critical thinking skills. This is not surprising in view of the fact that Japanese education does not seem to encourage

debate or the critical evaluation of reasoning (Davidson, 1995). Detailed consideration of the student answers themselves is beyond the scope of this study, but it is revealing to explore a little the kinds of errors consistently made by the participants. All of the subjects were taught to identify and use definition, illustration, and argumentation as rhetorical modes; however, this training apparently did not prepare them to recognize reasoning errors related to these modes of expression. For example, out of the 24 students who positively but incorrectly evaluated Paragraph 2, 24 gave as their justification the fact that Raywift provided a reason grounded in reality or else that he gave a concrete example. They missed the fact that both the reason and the example were irrelevant to his argument. Similarly, students accepted the definition-arguments in Paragraphs 1 and 7, even though the definitions offered by Raywift were inappropriate. For instance, he argues in Paragraph 7 that his opponents "don't know what 'safe' really means. Conditions are not safe if there's even the slightest possible chance for an accident" [italics added] (Ennis & Weir, 1985, p. 13). Only 2 students found fault with this impractical definition of the concept of safety; the others credited him with giving a clear definition. Likewise, 25 of 36 students accepted the false analogy used in Paragraph 1. Though the treatment group members fared better on some paragraphs and in their overall scores, these common tendencies seem to point to a general need for critical thinking training among these particular Japanese EFL students that perhaps is not being addressed adequately by practice in English rhetorical modes or content-based study. It is even possible that student exposure to rhetorical modes such as definition, illustration, and argumentation may only predispose them to accept weak ideas simply because they are presented in the proper rhetorical format. Without concurrent attention to reasoning fallacies and the pitfalls related to each mode, teachers may discover that for their EFL/ESL students, a little bit of knowledge of rhetorical modes is a dangerous thing. Such students may one day find themselves struggling with the reasoning tasks required in an English academic setting, regardless of their general English language proficiency or familiarity with English modes of expression.

This is only a limited quasi-experimental pilot study, and more research of a similar type needs to be done to substantiate these tentative conclusions. Larger student samples are needed. Also, it would be helpful if a translated version of the test could be administered to groups of similar Japanese students to remove completely the possibility that English language deficiencies may to some extent account for the lower scores. For cultural and linguistic reasons, however, such a translated test may be difficult to make and administer. Students in mixed-nationality EFL/ESL programs in other cultural settings could also provide interesting and relevant data about critical thinking abilities and the possibility of developing and testing them in English language programs, since English language-learning problems related to thinking are not confined to Japan. English instructors in other places have noted reasoning weaknesses similar to the ones we have found (Sherman, 1992; Matthews, 1994). Furthermore, it would be informative to experiment with other standard tests of critical thinking in EFL/ESL programs. Finally, it is worth exploring the question of whether training in critical thinking can improve general English language proficiency, especially in writing and reading. Nevertheless, we hope to see the Ennis-Weir test applied by others in studies bearing some similarity to ours. This relatively unexplored area invites further inquiry.

References

Al-arishi, A. (1994). Practical ways to promote reflection in the EFL/ESL classroom. English Teaching Forum, 32 (2), 2-5.

Babbie, E. (1983). The practice of social research (3rd ed.). Belmont, CA: Wadsworth.

Beyer, B. (1991). Practical strategies for the direct teaching of thinking skills. In A. Costa (Ed.), Developing minds: A resource book for teaching thinking: Vol. 1 (2nd ed., pp. 274-279). Alexandria: Association for Supervision and Curriculum Development.

Carrell, P. (1987). Content and formal schemata in ESL reading. TESOL Quarterly, 21 (3), 461-481.

Chaffee, J. (1990). Thinking critically. New York: Houghton-Mifflin.

Chamot, A. (1995). Creating a community of thinkers in the ESL/EFL classroom. TESOL Matters, 5, (5), 1, 16.

Chance (1986). Thinking in the classroom: A survey of programs. New York: Teacher's College Press, Columbia University.

Damer, E. (1995). Attacking faulty reasoning: A practical guide to fallacy-free arguments. Belmont: Wadsworth.

Davidson, B. (1994). Critical thinking: A perspective and prescriptions for language teachers. The Language Teacher, 18 (4), 20-26.

Davidson, B. (1995). Critical thinking education faces the challenge of Japan. Inquiry: Critical thinking across the disciplines, 14 (3), 41-53.

Engel, M. (1994). With good reason: An introduction to informal fallacies. New York: St. Martin's Press.

Ennis, R., Millman, J., & Tomko, T. (1985). Cornell Critical Thinking Tests Level X & Level Z--manual. Pacific Grove: Critical Thinking Press and Software.

Ennis, R. & Weir, E. (1985). The Ennis-Weir Critical Thinking Essay Test. Pacific Grove: Critical Thinking Press and Software.

Halpern, D. (1993). Assessing the effectiveness of critical-thinking instruction. JGE: The Journal of General Education, 42 (4), 238-254.

Hatcher, D. (1994, August). Grading the Ennis-Weir Critical Thinking Essay Test. Paper presented at the Fourteenth International Conference on Critical Thinking and Educational Reform, Rohnert Park, CA.

Hatcher, D. (1996). Combining critical thinking and written composition: The whole is greater than the sum of the parts. Inquiry: Critical thinking across the disciplines, 15 (2), 20-36.

Lippman, M. (1991). Thinking in education. Cambridge: Cambridge University Press.

Matthews, C. (1994, March). Using news stories to develop thinking/reasoning skills. Paper presented at the TESOL annual convention, Baltimore, Maryland.

Norris, S. & Ennis, R. (1989). Evaluating critical thinking. Pacific Grove: Critical Thinking Press and Software.

Paul, R. W. (1992). Critical thinking: What every person needs to survive in a rapidly changing world. Santa Rosa, CA: Foundation for Critical Thinking.

Scriven, M. (1976). Reasoning. New York: McGraw-Hill.

Sherman, J. (1992). Your own thoughts in your own words. ELT Journal, 46 (2), 190-198.

Siegel, H. (1988). Educating reason: Rationality, critical thinking, and education. New York: Routledge Press.

Snow, M. & Brinton, D. (1988). Content-based language instruction: Investigating the effectiveness of the adjunct model. TESOL Quarterly, 22 (4), 533-574.

Tarvin, W. & Al-arishi, A. (1991). Rethinking communicative language teaching: Reflection and the EFL classroom. TESOL Quarterly, 25 (1), 9-27.

Walsh, D. & Paul, R. (n.d.). The goal of critical thinking: From educational ideal to educational reality. American Federation of Teachers Educational Issues Department.

Table 1

Critical Thinking Skills Addressed on the Ennis-Weir Test

Paragraph	Skill
1.	Noticing misuse of analogy and/or shift in meaning
2.	Recognizing irrelevant reasoning
3.*	Recognizing relevant reasoning
4.	Recognizing circularity and/or the lack of a reason
5.	Recognizing defective reasoning
6.	Recognizing insufficient sampling
7.	Recognizing equivocation and/or the use of an arbitrary definition
8.*	Evaluating the credibility of expert testimony

*Paragraphs that exhibit sound reasoning

Table 2

Group Comparative Scores on the Ennis-Weir Test

Score Range	No. of Students Control Group	No. of Students Treatment Group
-4.0 to 0.0	9	1
1.0 to 2.0	6	1
2.5 to 6.5	4	5
7.0 to 13.5	0	10
TOTAL STUDENTS	19	17
Mean	0.6	6.6
Median	1.0	7.5
Mode	-1.5	3.0

Table 3

Student Scores on Ennis-Weir Test by English Level

Score Range	A-B Students	C Students	D-E Students
-4.0 to 0.0	4	4	2
1.0 to 6.5	4	8	4
7.0 to 13.5	4	2	4
TOTAL	12	14	10
Mean	4.3	1.8	4.7
Median	3.8	1.5	4.3
Mode	10.5	1.0	3.0

Biographical Note

Bruce W. Davidson teaches English at Hokusei Gakuen University in the Department of Social Welfare and does research related to critical thinking education. Rodney L. Dunham teaches in the English Department at Tezukayama College and also does research in mass communication.

Correspondence concerning this article should be addressed to Bruce W. Davidson, who is now at Hokusei Gakuen University, 3-1, Nishi 2 chome, Oyachi, Atsubetsu-ku, Sapporo, Japan 004. Electronic mail may be sent via Internet to z00134@hokusei.ac.jp.

Acknowledgment

This substance of this paper was presented on November 5, 1995 at the Twenty-First Annual JALT (Japan Association for Language Teaching) International Conference held in Nagoya, Japan. It will also be presented at the Sixteenth International Conference on Critical Thinking and Educational Reform held in Rohnert Park, CA, July 28-31, 1996.